Stappenplan IZB dashboard

# Overzicht

Dit dashboard is ontwikkeld bij GGD Noord- en Oost-Gelderland als pilot voor een schaalbare datastructuur die bovenregionaal ingezet kan worden. Het belangrijkste element is daarom het datamodel; het meegeleverde dashboard (in Power BI Desktop) is slechts ter illustratie. Het dashboard wat in productie genomen is (in Power BI Cloud), is het natuurlijke vervolg van dit concept. Deze kan echter niet makkelijk gedeeld worden. Om dubbel werk te voorkomen en uitwisseling te bevorderen wordt de achterliggende structuur in deze documenten toegelicht.

Het gedeelde dashboard kan vrij eenvoudig binnen een andere GGD worden opgezet. Enige vereisten hiervoor zijn een SQL server en R. Ondersteuning van de afdeling IT is hierin waarschijnlijk nodig, voornamelijk bij het configureren van de server. Hieronder het stappenplan globaal weergegeven. De uitwerking kunt u vinden in onderstaande alinea’s.

1. Opzetten datastructuur
2. Invullen medische database
3. Configureren R-script
4. Inladen data HPZone
5. Opbouwen en uitbreiden dashboard
6. Gebruik en onderhoud

# Opzetten datastructuur

De datastructuur is kant-en-klaar, en kan direct worden toegepast op een Microsoft SQL Server. Voor andere vormen van SQL is waarschijnlijk een verandering in syntax nodig. (Let bijvoorbeeld op het gebruik van nvarchar.) Voor deze handleiding wordt uitgegaan van Microsoft SQL Server. Indien er geen organisatieserver beschikbaar is kan er voor een proof-of-concept ook een lokale versie gratis worden gedownload bij Microsoft (https://www.microsoft.com/en-us/sql-server/sql-server-downloads).

In de map SQL scripts staan de gebruikte tabellen beschreven. Deze bestanden kunnen direct uitgevoerd worden, bij voorkeur in een nieuwe database. Voor gebruiksgemak zijn alle tabellen ook samengevoegd in één bestand (alle\_tabellen\_samen.sql). **Let op:** Voer het bestand uit tot regel 200. Vul daarna eerst de medische database. (Zie volgende sectie.) Voer daarna pas de regels na 200 uit.

In de datastructuur wordt meerdere keren verwezen naar groepen. Deze zijn toegevoegd vanwege het gebrek aan consistentie binnen HPZone. Er kunnen verschillende diagnoses gebruikt worden bij één specifieke infectieziekte. Zo zijn er bijvoorbeeld drie soorten hepatitis B, die bij het weergeven in een dashboard waarschijnlijk als één collectie weergegeven moeten worden. De tabel *infectieziekte* beschrijft deze groepen, en een overzicht van de groepering is beschikbaar als losse Excelsheet.

# Invullen medische database

HPZone bevat een lijst met ziekteverwekkers (agents), infecties, diagnoses, en meldingsplichtige ziektes (deze laatste is niet te exporteren), en daarnaast een aantal categorieën om details aan te geven, zoals een scenario, context, of onderwerp. Deze zijn in de datastructuur herkenbaar door de prefix *hpz\_*. De eerste elementen zijn landelijk gelijk, de laatste elementen kunnen per GGD verschillen. Daarom worden deze dynamisch toegevoegd vanuit de exports van HPZone.

Laad de meegeleverde, of eigen (HPZone > Reference > Clinical), bestanden uit de map export in de hierboven gemaakte database. Let hierbij vooral bij de onderwerpen, zoals gebruikt bij telefonische vragen (enquiries), op het indelen in groepen, indien gewenst. Het script import\_datadump.R kan gebruikt worden om de meegeleverde CSV’s snel in te laden. **Let op:** Voeg nu de constraints toe. Dit zijn regels 200 t/m het einde van alle\_tabellen\_samen.sql.

# Configureren R-script

Het inladen vanuit HPZone kan via 2 R-scripts. Eén script is bedoeld om gebruikt te worden met SQL Server Integration Services (preparatie\_HPZone\_export.R), en zet de HPZone-bestanden om in een CSV-formaat wat deze software prettig vindt. In dit script hoeft alleen de uitvoerlocatie opgegeven te worden. Het andere script (data HPZone API.R) maakt zelf een verbinding met de server en voegt de data toe/aan. Hiervoor is een API key nodig, welke waarschijnlijk per mail is ontvangen vanuit Prolocation. Stel voor de eerste uitvoering de jaren (bovenin het script) in op de gewenste periode, en vervang deze daarna weer door het codewoord “cur”. N.B.: De gebruikte packages kunnen gebruik maken van de huidige gebruiker binnen Windows, dus beveiliging op persoonsniveau is mogelijk. (Zie ook de documentatie: <https://cran.r-project.org/web/packages/odbc/index.html>)

# Inladen data HPZone

De data wordt automatisch opgehaald met de API. Het script kan daardoor simpelweg met enige regelmaat worden uitgevoerd, waarna alle data weer bijgewerkt is. Dagelijkse verversing is hierbij ruim voldoende.

# Opbouwen en uitbreiden dashboard

Een eerste conceptversie van het dashboard is meegeleverd als Power BI Desktop bestand (.pbit). Hiervoor is Power BI Desktop vereist, welke gratis gedownload kan worden in de Microsoft Store (geen administratorrechten vereist). Dit voorbeeld kan direct functioneel gemaakt worden door het invoeren van de juiste database, of als voorbeeld voor een eigen opbouw. Verder kan de datastructuur ook simpel ingeladen worden in een eigen softwarepakket (Qliksense, Swing, enz.); gebruik hiervoor de weergavetabellen, herkenbaar aan de prefix *vw\_*. Het is niet aan te raden om de brontabellen zelf te gebruiken, aangezien er dan een datamodel opgebouwd moet worden.

# Gebruik en onderhoud

De dataset kan gebruikt worden voor een dashboard, maar ook voor verdere analyse of onderzoek. Aanroepen vanuit R is bijzonder simpel, met de packages *DBI* en *odbc*. Hier kunnen makkelijk queries worden uitgevoerd. Zo kunnen de uitbraken (situations) van de laatste 6 weken opgevraagd worden met:

uitbraken = dbGetQuery(connection, “SELECT \* FROM vw\_situations WHERE datum >= (NOW – INTERVAL ‘6 weeks’”)

Het bijgevoegde script voor de analyse van gevallen in de laatste 6 weken maakt hier veelvuldig gebruik van: er worden automatisch grafieken gemaakt van de incidentie van de ziekten van meldingen in de laatste 6 weken. Hierdoor is een afwijking van het gebruikelijke patroon makkelijk te visualiseren.

Qua onderhoud hoeft er weinig te gebeuren. Alleen de data moet bijgewerkt worden, momenteel nog handmatig. Dit kan zo vaak als gewenst; de scripts om data bij te werken zijn niet bijzonder zwaar.